RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/930, 440 c

Source: IFW16

Date Processed by STIC: 12-14-04

ENTERED

```
        RAW SEQUENCE LISTING              DATE: 12/14/2004
        PATENT APPLICATION:  US/09/930,440C    TIME: 14:00:39

        Input Set : A:\03940077pa.txt
        Output Set: N:\CRF4\12142004\I930440C.raw

   3 <110> APPLICANT: Betenbaugh, Michael J.
   4        Lawrence, Shawn J.
   5        Lee, Yuan C.
   6        Coleman, Timothy A.
   8 <120> TITLE OF INVENTION: Engineering Intracellular Sialylation Pathways
  10 <130> FILE REFERENCE: 03940077pa
  12 <140> CURRENT APPLICATION NUMBER: 09/930,440C
  13 <141> CURRENT FILING DATE: 2001-08-16
  15 <150> PRIOR APPLICATION NUMBER: US 60/122,582
  16 <151> PRIOR FILING DATE: 1999-03-02
  18 <150> PRIOR APPLICATION NUMBER: US 60/169,624
  19 <151> PRIOR FILING DATE: 1999-12-08
  21 <150> PRIOR APPLICATION NUMBER: US 60/227,579
  22 <151> PRIOR FILING DATE: 2000-08-25
  24 <150> PRIOR APPLICATION NUMBER: US 09/516,793
  25 <151> PRIOR FILING DATE: 2000-03-01
  27 <160> NUMBER OF SEQ ID NOS: 18
  29 <170> SOFTWARE: PatentIn version 3.2
  31 <210> SEQ ID NO: 1
  32 <211> LENGTH: 1429
  33 <212> TYPE: DNA
  34 <213> ORGANISM: Homo sapiens
  36 <400> SEQUENCE: 1
  37 atggccttcc caaagaagaa acttcagggt cttgtggctg caaccatcac gccaatgact      60
  39 gagaatggag aaatcaactt ttcagtaatt ggtcagtatg tggattatct tgtgaaagaa     120
  41 cagggagtga agaacatttt tgtgaatggc acaacaggag aaggcctgtc cctgagcgtc     180
  43 tcagagcgtc gccaggttgc agaggagtgg gtgacaaaag ggaaggacaa gctggatcag     240
  45 gtgataattc acgtaggagc actgagcttg aaggagtcac aggaactggc ccaacatgca     300
  47 gcagaaatag gagctgatgg catcgctgtc attgcaccgt cttcctcaa gccatggacc      360
  49 aaagatatcc tgattaattt cctaaaggaa gtggctgctg ccgcccctgc cctgccattt     420
  51 tattactatc acattcctgc cttgacaggg gtaaagattc gtgctgagga gttgttggat     480
  53 gggattctgg ataagatccc caccttccaa gggctgaaat tcagtgatac agatctctta     540
  55 gacttcgggc aatgtgttga tcagaatcgc cagcaacagt ttgctttcct ttttggggtg     600
  57 gatgagcaac tgttgagtgc tctggtgatg ggagcaactg gagcagtggg cagttttgta     660
  59 tccagagatt tatcaacttt gttgtcaaac taggttttgg agtgtcacag accaaagcca     720
  61 tcatgactct ggtctctggg attccaatgg gcccaccccg gcttccactg cagaaagcct     780
  63 ccagggagtt tactgatagt gctgaagcta aactgaagag cctggatttc ctttctttca     840
  65 ctgatttaaa ggatggaaac ttggaagctg gtagctagtg cctctctatc aaatcagggt     900
  67 ttgcaccttg agacataatc taccttaaat agtgcatttt tttctcaggg aattttagat     960
  69 gaacttgaat aaactctcct agcaaatgaa atctcacaat aagcattgag gtacctttg     1020
  71 tgagccttaa aaagtcttat tttgtgaagg ggcaaaaact ctaggagtca caactctcag    1080
  73 tcattcattt cacagatttt tttgtggaga aatttctgtt tatatggatg aaatggaatc    1140
  75 aagaggaaaa ttgtaattga ttaattccat ctgtctttag gagctctcat tatctcggtc    1200
```

RAW SEQUENCE LISTING   DATE: 12/14/2004
PATENT APPLICATION: US/09/930,440C   TIME: 14:00:40

Input Set : A:\03940077pa.txt
Output Set: N:\CRF4\12142004\I930440C.raw

```
77 tctggttcct aatcctattt taaagttgtc taattttaaa ccactataat atgtcttcat    1260
79 tttaataaat attcatttgg aatctaggaa aactctgagc tactgcattt aggcaggcac    1320
81 tttaatacca aactgtaaca tgtctcaact gtatacaact caaaatacac cagctcattt    1380
83 ggctgctcag tctaactcta gaatggatgc ttttgaattc atttcgatg                1429
86 <210> SEQ ID NO: 2
87 <211> LENGTH: 304
88 <212> TYPE: PRT
89 <213> ORGANISM: Homo sapiens
91 <400> SEQUENCE: 2
93 Met Ala Phe Pro Lys Lys Lys Leu Gln Gly Leu Val Ala Ala Thr Ile
94 1               5                   10                  15
97 Thr Pro Met Thr Glu Asn Gly Glu Ile Asn Phe Ser Val Ile Gly Gln
98              20                  25                  30
101 Tyr Val Asp Tyr Leu Val Lys Glu Gln Gly Val Lys Asn Ile Phe Val
102             35                  40                  45
105 Asn Gly Thr Thr Gly Glu Gly Leu Ser Leu Ser Val Ser Glu Arg Arg
106         50                  55                  60
109 Gln Val Ala Glu Glu Trp Val Thr Lys Gly Lys Asp Lys Leu Asp Gln
110 65                  70                  75                  80
113 Val Ile Ile His Val Gly Ala Leu Ser Leu Lys Glu Ser Gln Glu Leu
114                 85                  90                  95
117 Ala Gln His Ala Ala Glu Ile Gly Ala Asp Gly Ile Ala Val Ile Ala
118                 100                 105                 110
121 Pro Phe Phe Leu Lys Pro Trp Thr Lys Asp Ile Leu Ile Asn Phe Leu
122             115                 120                 125
125 Lys Glu Val Ala Ala Ala Ala Pro Ala Leu Pro Phe Tyr Tyr Tyr His
126     130                 135                 140
129 Ile Pro Ala Leu Thr Gly Val Lys Ile Arg Ala Glu Glu Leu Leu Asp
130 145                 150                 155                 160
133 Gly Ile Leu Asp Lys Ile Pro Thr Phe Gln Gly Leu Lys Phe Ser Asp
134                 165                 170                 175
137 Thr Asp Leu Leu Asp Phe Gly Gln Cys Val Asp Gln Asn Arg Gln Gln
138                 180                 185                 190
141 Gln Phe Ala Phe Leu Phe Gly Val Asp Glu Gln Leu Leu Ser Ala Leu
142         195                 200                 205
145 Val Met Gly Ala Thr Gly Ala Val Gly Ser Phe Val Ser Arg Asp Leu
146     210                 215                 220
149 Ser Thr Leu Leu Ser Asn Val Leu Glu Cys His Arg Pro Lys Pro Ser
150 225                 230                 235                 240
153 Leu Trp Ser Leu Gly Phe Gln Trp Ala His Pro Gly Phe His Cys Arg
154                 245                 250                 255
157 Lys Pro Pro Gly Ser Leu Leu Ile Val Leu Lys Leu Asn Arg Ala Trp
158                 260                 265                 270
161 Ile Ser Phe Leu Ser Leu Ile Arg Met Glu Thr Trp Lys Leu Val Ala
162             275                 280                 285
165 Ser Ala Ser Leu Ser Asn Gln Gly Phe Ala Pro Leu Arg His Asn Leu
166     290                 295                 300
169 <210> SEQ ID NO: 3
170 <211> LENGTH: 1305
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/930,440C

DATE: 12/14/2004
TIME: 14:00:40

Input Set : **A:\03940077pa.txt**
Output Set: **N:\CRF4\12142004\I930440C.raw**

```
171 <212> TYPE: DNA
172 <213> ORGANISM: Homo sapiens
174 <400> SEQUENCE: 3
175 atggactcgg tggagaaggg ggccgccacc tccgtctcca acccgcgggg gcgaccgtcc      60
177 cggggccggc cgccgaagct gcagcgcaac tctcgcggcg gccagggccg aggtgtggag     120
179 aagcccccgc acctggcagc cctaattctg gcccgggggg gcagcaaagg catccccctg     180
181 aagaacatta agcacctggc gggggtcccg ctcattggct gggtcctgcg tgcggccctg     240
183 gattcagggg ccttccagag tgtatggggtt tcgacagacc atgatgaaat tgagaatgtg     300
185 gccaaacaat ttggtgcaca agttcatcga agaagttctg aagtttcaaa agacagctct     360
187 acctcactag atgccatcat agaatttctt aattatyata atgaggktga cattgtagga     420
189 aatattcaag ctacttctyc atgtttacat cctactgatc ttcaaaaagt tgcagaaatg     480
191 attcgagaag aaggatatga ttctgktttc tctgttgtga gacgccatca gtttcgatgg     540
193 agtgaaattc agaaaggagt tcgtgaagtg accgaacctc tgaatttaaa tccagctaaa     600
195 cggcctcgtc gacaagactg ggatggagaa ttatatgaaa atggctcatt ttattttgct     660
197 aaaagacatt tgatagagat gggttacttg cagggtggaa aaatggcata ctacgaaatg     720
199 cgagctgaac atagtgtgga tatagatgtg gatattgatt ggcctattgc agagcaaaga     780
201 gtattaagat atggctattt tggcaaagag aagcttaagg aaataaaaact tttggttttgc     840
203 aatattgatg gatgtctcac caatggccac atttatgtat caggagacca aaaagaaata     900
205 atatcttatg atgtaaaaga tgctattggg ataagtttat taaagaaaag tggtattgag     960
207 gtgaggctaa tctcagaaag ggcctgttca aagcagacgc tgtcttcttt aaaactggat    1020
209 tgcaaaatgg aagtcagtgt atcagacaag ctagcagttg tagatgaatg gagaaaagaa    1080
211 atgggcctgt gctggaaaga gtggcatat cttggaaatg aagtgtctga tgaagagtgc    1140
213 ttgaagagag tgggcctaag tggcgctcct gctgatgcct gttcctacgc ccagaaggct    1200
215 gttggataca tttgcaaatg taatggtggc cgtggtgcca tccgagaatt tgcagagcac    1260
217 atttgcctac taatggaaaa agttaataat tcatgccaaa aatag                    1305
220 <210> SEQ ID NO: 4
221 <211> LENGTH: 434
222 <212> TYPE: PRT
223 <213> ORGANISM: Homo sapiens
226 <220> FEATURE:
227 <221> NAME/KEY: misc_feature
228 <222> LOCATION: (133)..(133)
229 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
231 <220> FEATURE:
232 <221> NAME/KEY: misc_feature
233 <222> LOCATION: (136)..(136)
234 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
236 <220> FEATURE:
237 <221> NAME/KEY: misc_feature
238 <222> LOCATION: (147)..(147)
239 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
241 <220> FEATURE:
242 <221> NAME/KEY: misc_feature
243 <222> LOCATION: (169)..(169)
244 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
246 <400> SEQUENCE: 4
248 Met Asp Ser Val Glu Lys Gly Ala Ala Thr Ser Val Ser Asn Pro Arg
249 1               5                   10                  15
252 Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Asn Ser Arg
```

RAW SEQUENCE LISTING   DATE: 12/14/2004
PATENT APPLICATION: US/09/930,440C   TIME: 14:00:40

Input Set : A:\03940077pa.txt
Output Set: N:\CRF4\12142004\I930440C.raw

```
   253            20               25               30
   256 Gly Gly Gln Gly Arg Gly Val Glu Lys Pro Pro His Leu Ala Ala Leu
   257         35               40               45
   260 Ile Leu Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu Lys Asn Ile Lys
   261      50               55               60
   264 His Leu Ala Gly Val Pro Leu Ile Gly Trp Val Leu Arg Ala Ala Leu
   265 65               70               75               80
   268 Asp Ser Gly Ala Phe Gln Ser Val Trp Val Ser Thr Asp His Asp Glu
   269              85               90               95
   272 Ile Glu Asn Val Ala Lys Gln Phe Gly Ala Gln Val His Arg Arg Ser
   273              100              105              110
   276 Ser Glu Val Ser Lys Asp Ser Ser Thr Ser Leu Asp Ala Ile Ile Glu
   277         115              120              125
W--> 280 Phe Leu Asn Tyr Xaa Asn Glu Xaa Asp Ile Val Gly Asn Ile Gln Ala
   281      130              135              140
W--> 284 Thr Ser Xaa Cys Leu His Pro Thr Asp Leu Gln Lys Val Ala Glu Met
   285 145              150              155              160
W--> 288 Ile Arg Glu Glu Gly Tyr Asp Ser Xaa Phe Ser Val Val Arg Arg His
   289              165              170              175
   292 Gln Phe Arg Trp Ser Glu Ile Gln Lys Gly Val Arg Glu Val Thr Glu
   293         180              185              190
   296 Pro Leu Asn Leu Asn Pro Ala Lys Arg Pro Arg Arg Gln Asp Trp Asp
   297         195              200              205
   300 Gly Glu Leu Tyr Glu Asn Gly Ser Phe Tyr Phe Ala Lys Arg His Leu
   301      210              215              220
   304 Ile Glu Met Gly Tyr Leu Gln Gly Gly Lys Met Ala Tyr Tyr Glu Met
   305 225              230              235              240
   308 Arg Ala Glu His Ser Val Asp Ile Asp Val Asp Ile Asp Trp Pro Ile
   309              245              250              255
   312 Ala Glu Gln Arg Val Leu Arg Tyr Gly Tyr Phe Gly Lys Glu Lys Leu
   313         260              265              270
   316 Lys Glu Ile Lys Leu Leu Val Cys Asn Ile Asp Gly Cys Leu Thr Asn
   317         275              280              285
   320 Gly His Ile Tyr Val Ser Gly Asp Gln Lys Glu Ile Ile Ser Tyr Asp
   321      290              295              300
   324 Val Lys Asp Ala Ile Gly Ile Ser Leu Leu Lys Lys Ser Gly Ile Glu
   325 305              310              315              320
   328 Val Arg Leu Ile Ser Glu Arg Ala Cys Ser Lys Gln Thr Leu Ser Ser
   329              325              330              335
   332 Leu Lys Leu Asp Cys Lys Met Glu Val Ser Val Ser Asp Lys Leu Ala
   333         340              345              350
   336 Val Val Asp Glu Trp Arg Lys Glu Met Gly Leu Cys Trp Lys Glu Val
   337         355              360              365
   340 Ala Tyr Leu Gly Asn Glu Val Ser Asp Glu Glu Cys Leu Lys Arg Val
   341      370              375              380
   344 Gly Leu Ser Gly Ala Pro Ala Asp Ala Cys Ser Tyr Ala Gln Lys Ala
   345 385              390              395              400
   348 Val Gly Tyr Ile Cys Lys Cys Asn Gly Gly Arg Gly Ala Ile Arg Glu
   349              405              410              415
```

```
352 Phe Ala Glu His Ile Cys Leu Leu Met Glu Lys Val Asn Asn Ser Cys
353             420             425             430
356 Gln Lys
360 <210> SEQ ID NO: 5
361 <211> LENGTH: 1080
362 <212> TYPE: DNA
363 <213> ORGANISM: Homo sapiens
365 <400> SEQUENCE: 5
366 atgccgctgg agctggagct gtgtcccggg cgctgggtgg gcgggcaaca cccgtgcttc      60
368 atcattgccg agatcggcca gaaccaccag ggcgacctgg acgtagccaa gcgcatgatc     120
370 cgcatggcca aggagtgtgg ggctgattgt gccaagttcc agaagagtga gctagaattc     180
372 aagtttaatc ggaaagcctt ggagaggcca tacacctcga agcattcctg ggggaagacg     240
374 tacggggagc acaaacgaca tctggagttc agccatgacc agtacaggga gctgcagagg     300
376 tacgccgagg aggttgggat cttcttcact gcctctggca tggatgagat ggcagttgaa     360
378 ttcctgcatg aactgaatgt tccatttttc aaagttggat ctggagacac taataatttt     420
380 ccttatctgg aaaagacagc caaaaaaggt cgcccaatgg tgatctccag tgggatgcag     480
382 tcaatggaca ccatgaagca gtttatcag atcgtgaagc ccctcaaccc caacttctgc     540
384 ttcttgcagt gtaccagcgc atacccgctc cagcctgagg acgtcaacct gcgggtcatc     600
386 tcggaatatc agaagctctt tcctgacatt cccataggt attctgggca tgaaacaggc     660
388 atagcgatat ctgtggccgc agtggctctg ggggccaagg tgttggaacg tcacataact     720
390 ttggacaaga cctggaaggg gagtgaccac tcggcctcgc tggagcctgg agaactggcc     780
392 gagctggtgc ggtcagtgcg tcttgtggag cgtgccctgg gctccccaac caagcagctg     840
394 ctgccctgtg agatggcctg caatgagaag ctgggcaagt ctgtggtggc caaagtgaaa     900
396 attccggaag gcaccattct aacaatggac atgctcaccg tgaaggtggg tgagcccaaa     960
398 gcctatcctc ctgaagacat ctttaatcta gtgggcaaga aggtcctggt cactgttgaa    1020
400 gaggatgaca ccatcatgga agaattggta gataatcatg gcaaaaaaat caagtcttaa    1080
403 <210> SEQ ID NO: 6
404 <211> LENGTH: 359
405 <212> TYPE: PRT
406 <213> ORGANISM: Homo sapiens
408 <400> SEQUENCE: 6
410 Met Pro Leu Glu Leu Glu Leu Cys Pro Gly Arg Trp Val Gly Gly Gln
411 1               5              10              15
414 His Pro Cys Phe Ile Ile Ala Glu Ile Gly Gln Asn His Gln Gly Asp
415             20              25              30
418 Leu Asp Val Ala Lys Arg Met Ile Arg Met Ala Lys Glu Cys Gly Ala
419         35              40              45
422 Asp Cys Ala Lys Phe Gln Lys Ser Glu Leu Glu Phe Lys Phe Asn Arg
423     50              55              60
426 Lys Ala Leu Glu Arg Pro Tyr Thr Ser Lys His Ser Trp Gly Lys Thr
427 65              70              75              80
430 Tyr Gly Glu His Lys Arg His Leu Glu Phe Ser His Asp Gln Tyr Arg
431             85              90              95
434 Glu Leu Gln Arg Tyr Ala Glu Glu Val Gly Ile Phe Phe Thr Ala Ser
435             100             105             110
438 Gly Met Asp Glu Met Ala Val Glu Phe Leu His Glu Leu Asn Val Pro
439         115             120             125
442 Phe Phe Lys Val Gly Ser Gly Asp Thr Asn Asn Phe Pro Tyr Leu Glu
443     130             135             140
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 12/14/2004
PATENT APPLICATION: **US/09/930,440C**   TIME: 14:00:41

Input Set : **A:\03940077pa.txt**
Output Set: **N:\CRF4\12142004\I930440C.raw**

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; Xaa Pos. 133,136,147,169
Seq#:9; N Pos. 1,3,6,12,13,15,18
Seq#:10; N Pos. 3,4,6,9,10,11,12,15,17,18
Seq#:11; N Pos. 3,4,6,9,10,11,12,15,17,18
Seq#:12; N Pos. 2,3,6,9,12,15,16,17,18

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:9,10,11,12,13,14,15,16

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/930,440C

DATE: 12/14/2004
TIME: 14:00:41

Input Set : **A:\03940077pa.txt**
Output Set: **N:\CRF4\12142004\I930440C.raw**

```
L:280 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:128
L:284 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:144
L:288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:160
L:686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:750 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
L:814 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:873 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
```